



ELSEVIER

Available online at [www.sciencedirect.com](http://www.sciencedirect.com)

SCIENCE @ DIRECT®

Global and Planetary Change 38 (2003) 31–53

GLOBAL AND PLANETARY  
CHANGE

[www.elsevier.com/locate/gloplacha](http://www.elsevier.com/locate/gloplacha)

## Simulation of high latitude hydrological processes in the Torne–Kalix basin: PILPS Phase 2(e) 2: Comparison of model results with observations

Bart Nijssen<sup>a,\*</sup>, Laura C. Bowling<sup>a</sup>, Dennis P. Lettenmaier<sup>a</sup>, Douglas B. Clark<sup>b</sup>,  
Mustapha El Maayar<sup>c</sup>, Richard Essery<sup>d</sup>, Sven Goers<sup>e</sup>, Yeugeniy M. Gusev<sup>f</sup>,  
Florence Habets<sup>g</sup>, Bart van den Hurk<sup>h</sup>, Jiming Jin<sup>i</sup>, Daniel Kahan<sup>j</sup>, Dag Lohmann<sup>k</sup>,  
Xieyao Ma<sup>l</sup>, Sarith Mahanama<sup>m</sup>, David Mocko<sup>n</sup>, Olga Nasonova<sup>f</sup>, Guo-Yue Niu<sup>i</sup>,  
Patrick Samuelsson<sup>o,p</sup>, Andrey B. Shmakin<sup>q</sup>, Kumiko Takata<sup>l</sup>, Diana Versegny<sup>r</sup>,  
Pedro Viterbo<sup>s</sup>, Youlang Xia<sup>t,1</sup>, Yongkang Xue<sup>j</sup>, Zong-Liang Yang<sup>u</sup>

<sup>a</sup>Department of Civil and Environmental Engineering, University of Washington, Seattle, WA, USA

<sup>b</sup>Centre for Ecology and Hydrology, Wallingford, UK

<sup>c</sup>Canadian Forest Service, Northern Forestry Centre, Edmonton, AB, Canada

<sup>d</sup>Met Office, UK

<sup>e</sup>GKSS Research Center, Geesthacht, Germany

<sup>f</sup>Institute of Water Problems, Russian Academy of Sciences, Russian Federation

<sup>g</sup>National Center for Atmospheric Research/CNRM, Toulouse, France

<sup>h</sup>Royal Netherlands Meteorological Institute, De Bilt, The Netherlands

<sup>i</sup>Department of Hydrology and Water Resources, University of Arizona, Tucson, AZ, USA

<sup>j</sup>University of California Los Angeles, Torrance, CA, USA

<sup>k</sup>NCEP Environmental Modeling Center (NOAA/NWS), USA

<sup>l</sup>Frontier Research System for Global Change, Japan

<sup>m</sup>Hydrological Sciences Branch, NASA Goddard Space Flight Center and GEST/University of Maryland Baltimore County, USA

<sup>n</sup>NASA Goddard Space Flight Center and SAIC/General Sciences Operation, USA

<sup>o</sup>Swedish Meteorological and Hydrological Institute, Sweden

<sup>p</sup>Department of Meteorology, Stockholm University, Stockholm, Sweden

<sup>q</sup>Institute of Geography, Russian Academy of Sciences, Moscow, Russian Federation

<sup>r</sup>Climate Research Branch, Meteorological Service of Canada, Canada

<sup>s</sup>European Centre for Medium-Range Weather Forecasts, Reading, UK

<sup>t</sup>Macquarie University, North Ryde, NSW, Australia

<sup>u</sup>Department of Geological Sciences, University of Texas at Austin, Austin, TX, USA

Received 9 December 2001; received in revised form 21 March 2002; accepted 15 June 2002

\* Corresponding author. Present address: Departments of Civil Engineering and Engineering Mechanics/Hydrology and Water Resources, University of Arizona, JW Harshbarger 320A, P.O. Box 210011, Tucson, AZ 85721-0011, USA. Tel.: +1-520-626-1277; fax: +1-520-621-1422.  
E-mail address: [nijssen@u.arizona.edu](mailto:nijssen@u.arizona.edu) (B. Nijssen).

<sup>1</sup> Present address: Institute for Geophysics, University of Texas at Austin, Austin, TX, USA.

## Abstract

Model results from 21 land-surface schemes (LSSs) designed for use in numerical weather prediction and climate models are compared with each other and with observations in the context of the Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) Phase 2(e) model intercomparison experiment. This experiment focuses on simulations of land-surface water and energy fluxes in the 58,000-km<sup>2</sup> Torne and Kalix river systems in northern Scandinavia, during the period 1989–1998. All models participating in PILPS Phase 2(e) capture the broad dynamics of snowmelt and runoff, but large differences in snow accumulation and ablation, turbulent heat fluxes, and streamflow exist. The greatest among-model differences in energy and moisture fluxes in these high-latitude environments occur during the spring snowmelt period, reflecting different model parameterizations of snow processes. Differences in net radiation are governed by differences in the simulated radiative surface temperature during the winter months and by differences in surface albedo during the spring/early summer. Differences in net radiation are smallest during the late summer when snow is absent. Although simulated snow sublimation is small for most models, a few models show annual snow sublimation of about 100 mm. These differences in snow sublimation appear to be largely dependent on differences in snow surface roughness parameterizations. The models with high sublimation generally lose their snowpacks too early compared to observations and underpredict the annual runoff. Differences in runoff parameterizations are reflected in differences in daily runoff statistics. Although most models show a greater variability in daily streamflow than the observations, the models with the greatest variability (as much as double the observed variability), produce most of their runoff through fast response, surface runoff mechanisms. As a group, those models that took advantage of an opportunity to calibrate to selected small catchments and to transfer calibration results to the basin at large had a smaller bias and root mean squared error (RMSE) in daily streamflow simulations compared with the models that did not calibrate.

© 2003 Elsevier Science B.V. All rights reserved.

*Keywords:* Torne–Kalix basin; PILPS Phase 2(e); hydrological processes

## 1. Introduction

This paper is the second in a three paper series that describes the design, implementation, analysis, and results of Phase 2(e) of the Project for Intercomparison of Land-surface Parameterization Schemes (PILPS). The first paper describes the history, purpose, design, and implementation of the PILPS Phase 2(e) experiment and provides a summary discussion of results (Bowling et al., 2003a-this issue). The third paper further examines the sensitivities of the land-surface schemes (LSSs) to changes in environmental conditions (Bowling et al., 2003b-this issue). In the current paper, simulation results from each of the LSSs are compared with available observations and with each other, to assess the models' ability to adequately capture the controlling processes in boreal and alpine environments.

The family of PILPS-2 experiments involves off-line testing of LSSs. Off-line testing means that prescribed atmospheric conditions are used to drive the LSSs and that there is no mechanism for repre-

sentation of feedbacks from the land surface to the atmosphere. In the PILPS-2 series of experiments, observed atmospheric forcing data are used as much as possible (Henderson-Sellers et al., 1995). The objective of the PILPS Phase 2(e) experiment is to “evaluate the performance of uncoupled land-surface parameterizations in high latitudes, in a context that allows evaluation of their ability to capture key processes spatially” (Bowling et al., 2003a-this issue). The Torne and Kalix river systems in northern Scandinavia, which have a combined drainage area of about 58,000 km<sup>2</sup> (Fig. 1), were selected as the study area to take advantage of observations collected by the Swedish Hydrological and Meteorological Institute and the Finnish Meteorological Institute (World Climate Research Programme (WCRP), 1999).

The 21 PILPS Phase 2(e) participants were provided with atmospheric forcings for the period 1979–1998. The first 10 years were available for model initialization and spin-up. Submitted results and the subsequent analyses were limited to the second 10-year period, 1989–1998. Full details of the experimental design, the

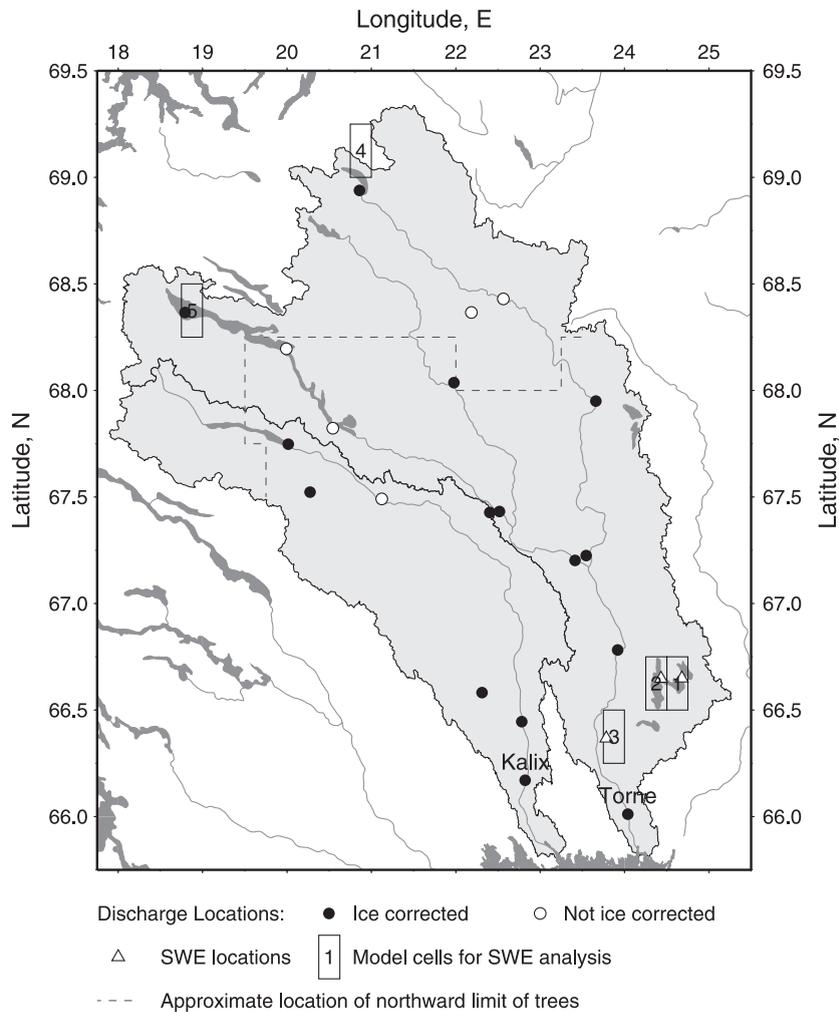


Fig. 1. Map of the Torne and Kalix river basins in northern Scandinavia. Superimposed are the discharge and snow locations discussed in the paper.

forcing data and the submitted results can be found in [Bowling et al. \(2003a-this issue\)](#).

One of the main difficulties in performing model studies at high northern latitudes is the lack of consistent, long-term, high-quality observations that can be used to evaluate model simulations. The Torne and Kalix River basins were selected for PILPS 2(e) because of their relatively dense surface meteorological network. Although the Torne and Kalix Rivers do not drain to the Arctic Ocean, their hydroclimatological characteristics are similar to those of many Arctic rivers. Furthermore, these river basins lie within the BALTEX domain (Baltic Sea Experiment), which is a

continental-scale study area of the World Climate Research Programme's GEWEX (Global Water and Energy Experiment) project. In addition to the model forcing data, discussed in [Bowling et al. \(2003a-this issue\)](#), a number of observations were collected in the Torne and Kalix River basins that can be used to study model performance. Even so, point observations of cold season processes are scarce in general and direct observations of the spatial variability of these processes are even scarcer. As a result, the comparison of simulated results with observations will largely focus on snow (extent, accumulation, and ablation) and streamflow.

After a brief discussion of water balance results (Section 2), the paper continues with a presentation and discussion of the simulations of snow spatial extent, snow accumulation, and snow ablation (Section 3). This is followed by a comparison of simulated and observed streamflow, both with respect to timing and magnitude (Section 4). Energy balance, turbulent energy fluxes, and net radiation results are evaluated in Section 5. The land-surface temperature and shortwave albedo and their effect on the energy balance during snowmelt periods are the topics in Section 6, which is followed by a discussion of the land-surface temperature and the sensible heat flux (Section 7) and the effect of roughness on the turbulent fluxes (Section 8).

## 2. Water balance

As discussed in [Bowling et al. \(2003a-this issue\)](#), model errors in the basin-wide, mean annual water balance were calculated as the difference between precipitation and the sum of total evaporation, surface, and subsurface runoff and the total storage change, including soil moisture, surface water, snow, and

canopy interception (see [Fig. 14 in Bowling et al., 2003a-this issue](#)). One model, RCA (B) (see [Table 1](#) for model labels), had a water balance error greater than the specified threshold of 3 mm/year and therefore failed the consistency criterion, which was adopted from the PILPS Phase 2(c) experiment ([Wood et al., 1998](#)). The basin-wide water balance error for RCA (B) peaks during snowmelt, with a residual error that is relatively constant throughout the year. Some of the models showed large positive (VISA (M)) or negative (SAST (N), MOSES (T), and MOSES-CEH (U)) net changes in moisture storage over the duration of the model run, which indicates a potential lack of equilibrium in the model simulations, despite the prescribed 10-year model spin-up period. In the case of MOSES and MOSES-CEH, evaporation from lakes is calculated, but a full water budget is not maintained for the lakes. To balance the basin water budget, lake evaporation was reported as a change in surface storage, and this term dominates the 10-year storage change. Recalculation of the storage change ignoring lake evaporation results in a  $-1.78$  mm/year change rather than a  $-15.05$  mm/year change for MOSES.

Table 1  
Model groups, names, and identifiers (see also [Table 6 in Bowling et al., 2003a-this issue](#))

Group	Model Name	Model Identifier
Group 1	SPONSOR	A
	RCA	B
Group 2	IHAS	C
	SEWAB	D
	ISBA	E
	NSIPP	F
	CLASS	G
	IBIS	H
	CHASM	I
Group 3	VIC	J
	MATSIRO	K
	HY-SSiB	L
	VISA	M
	SAST	N
	MECMWF	O
	NOAH	P
Group 4	SWAP	Q
	SSiB	R
	ECMWF	S
	MOSES	T
	MOSES-CEH	U

## 3. Snow

### 3.1. Snow spatial extent

Remotely sensed observations offer the only practical way to assess the evolution of snow extent over large areas. The Northern Hemisphere EASE-Grid Weekly Snow Cover and Sea Ice Extent product, obtained from the EOSDIS NSIDC Distributed Active Archive Center, combines snow cover and sea ice extent at weekly intervals from October 1978 through August 1995. The EASE-Grid product is derived from digitized versions of the NOAA-NESDIS Weekly Northern Hemisphere Snow Charts, which are manual interpretations of AVHRR, GOES, and other visible-band satellite data. The original, manually derived source snow charts have pixel sizes ranging from  $125 \times 125$  to  $205 \times 205$  km, much coarser than the  $25 \times 25$  km final EASE-Grid product ([National Snow and Ice Data Center, 1996](#)), which has a spatial resolution similar to the PILPS Phase 2(e) computational grid. The PILPS Phase 2(e) grid consists of  $218 \frac{1}{4}^\circ$  grid cells, which at  $67^\circ\text{N}$  corresponds to a grid

cell size of about 28 km east–west and 11 km north–south.

During the years in the study period for which images are available (1989–1995), the study area was completely covered by snow until at least the middle of April, and during some years the area remained snow covered until the 2nd week of May. Snow-free conditions first occur in the southern, low-lying part of the basin and progress northward over a period of 3 to 4 weeks. Snow cover persists longest in the northwest, which is the highest and wettest part of the study

area. The remote sensing product indicates that the northern half of the study area stays snow covered well into June during certain years, while in other years snow disappears for a brief period in May and then reappears in June. However, this reappearance of snow is not confirmed by ground-based snow depth observations in this area (Section 3.2), and is perhaps an artifact of the gridding and classification procedure used in the generation of the remote sensing product. This discrepancy between the remote sensing product and ground-based observations was especially large in

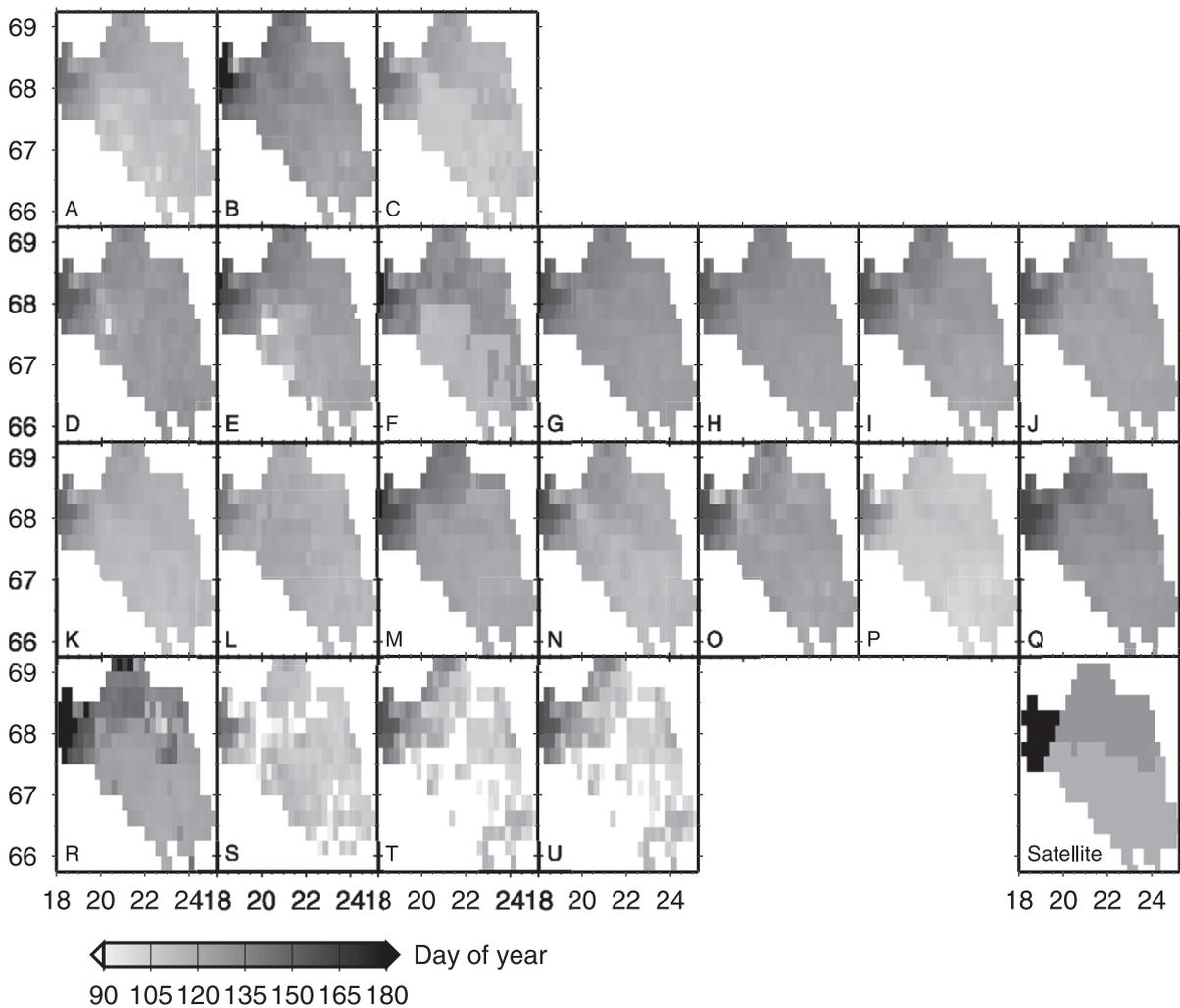


Fig. 2. Satellite-based (lower right) and simulated mean first snow-free day in the Torne and Kalix river basins for the period 1989–1993. The figure shows the advance of snowmelt from the south to the northwest. Darker colors indicate grid cells that are snow covered till later in the year. For model labels see Table 1.

1994 and 1995 and these years were excluded from the subsequent snow cover extent analysis.

Fig. 2 shows the mean simulated and satellite-observed advance of snowmelt over the Torne–Kalix River basins for the 1989–1993 period. It was constructed by determining the first snow-free day in each pixel for each year and then averaging over all the years. For analysis purposes, the first snow-free day in each year was taken to be the first day that the reported average snow water equivalent in a model grid cell was less than 10 mm. Although the selection of the 10-mm threshold is somewhat arbitrary, a non-zero threshold was selected because the satellite product will classify a pixel as snow free when there is still some patchy snow on the ground. Selection of a zero threshold would result in an artificial late bias in the model results compared to the observations. For the satellite-based gridded snow cover product, the first snow-free day was taken to be the first day that a pixel was classified as snow free. The low resolution of the original snow charts is reflected in the lack of detail in the satellite-observed progression of snowmelt. Note that the models in Fig. 2 (and in the following figures) are grouped in the same way as in Bowling et al. (2003a-this issue). Model identifiers (A–U) correspond to the model identifiers and names in Table 1 (see also Table 6 in Bowling et al., 2003a-this issue). Model groupings are based on the simulated mean annual net radiation and the mean annual latent heat flux (see Bowling et al. (2003a-this issue) for details).

All models capture the broad features of the advance of the first snow-free day from south to north with snow lingering in the northwest part of the study area (Fig. 2). However, the timing of the first snow-free day in the south varies by more than 6 weeks among models, with a few models about 4 weeks early and some models 2 to 3 weeks late compared with the satellite product. The models in groups 1 and 4, identified in Bowling et al. (2003a-this issue) as the models with the largest mean annual latent heat flux, generally lose most of their snow cover very early in the season. The models in the other two groups show no clear pattern, with some of the models losing most of their snow cover early (e.g., NOAH (P) in group 3) and some models keeping their snow cover until late in the spring (for example SEWAB (D) in group 2 and SWAP (Q) in group 3). Averaged over the entire study

period, instead of the period 1989–1993, the spatial patterns of snowmelt are similar, although the snow cover generally disappears somewhat later than during the first 5 years. Some of the models in groups 2 and 3 (SEWAB (D), ISBA (E) and MECMWF (O)) show single pixels where snow disappears early in the season.

### 3.2. Snow accumulation, melt, and sublimation

Daily snow depth observations for the period 1979–1998 were provided for eight locations in the Swedish part of the basin by the Swedish Meteorological and Hydrological Institute (SMHI). Snow was measured at 7:00 AM local time on a more or less flat surface where no drifts had formed. The reported snow depths represent the average value of at least five points a few meters apart (Dahlstrom, 1995). Snow depth observations were used to quality check the remote sensing product discussed in the previous section.

Observations of snow water equivalent were provided for three low elevation locations in the Finnish part of the basin by the Finnish Environment Institute (Fig. 1). The snow water equivalent was estimated from snow water equivalent measurements along transects taken on the 1st and 16th day of every month during the period 1979–1999. The snow measurements were taken over areas of 81, 107, and 347 km<sup>2</sup>. The observed snow water equivalent for the three sites is compared with the simulated snow water equivalent in the corresponding 1/4° grid cells for the first part of 1995 in Fig. 3. The simulated snow water equivalent for two additional grid cells, for which no observations were available, but which had different vegetation cover, are included as well (see Fig. 1 for the locations of the grid cells). The mean elevation, total amount of solid and liquid precipitation and mean annual temperature for the five sites are given in Table 2. The first three sites have similar land cover, elevation, and climate. The observed snow water equivalent is similar for cells 1, 2, and 3 and reaches a maximum at the end of March of about 180 mm. No observations are available after May 16. Most models show little among-site variation in their simulations of the three sites with observations, with similar maximum accumulations and a similar decline of the snowpack during melt. There are some excep-

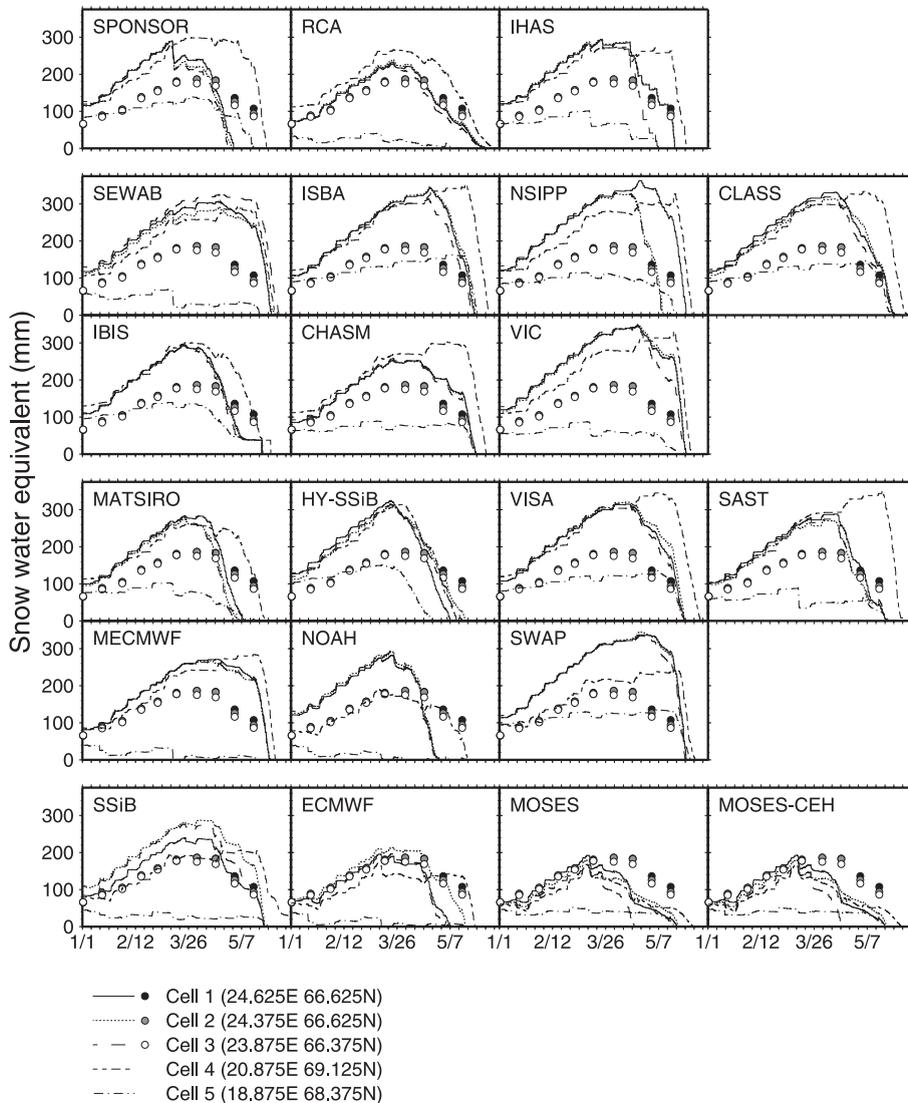


Fig. 3. Observed (dots) and simulated (lines) snow water equivalent for five locations during the first part of 1995 (no observations were available for cells 4 and 5).

tions, however. For example, in NSIPP (F) melt occurs about 2 1/2 weeks later in cell 1 than in cells 2 and 3. For ECMWF (S), melt in cell 2 occurs later than in cells 1 and 3, while for IHAS (C) melt occurs later in cells 1 and 2 than in cell 3. However, the among-model differences in maximum accumulation and timing of melt are much more pronounced (Fig. 3). Almost all models accumulate more snow than observed. Although there is a discrepancy

between the resolution of the model simulation and the local scale at which measurements are made, the close agreement between the observations from the three separate measurement sites suggests that the observations are fairly robust. Note that the reported values at each site represent an average of a number of point measurements along one or more snow transects. The only models for which the maximum accumulated amount is similar to that observed are the four

Table 2

Land cover, elevation, and atmospheric characteristics for the five sites in Figs. 3 and 4

	Cell 1 (24.625°E, 66.625°N)	Cell 2 (24.375°E, 66.625°N)	Cell 3 (23.875°E, 66.375°N)	Cell 4 (20.875°E, 69.125°N)	Cell 5 (18.875°E, 68.375°N)
Land cover (%)					
Water	23.5	32.5		6.0	35.1
Evergreen forest	1.2	3.2			
Woodland	32.0	46.1	52.9		
Wooded grassland	43.3	18.2	47.1		10.7
Closed shrubs				20.4	20.7
Open shrubs				63.1	29.3
Grassland					4.2
Bare ground				10.5	
Elevation (m M.S.L.)	127	120	120	765	633
Atmospheric characteristics (mean annual)					
Snowfall (mm)	428	429	457	375	306
Rainfall (mm)	265	264	262	368	256
Air temperature (°C)	0.8	0.9	1.1	−2.3	−1.0

models in group 4 (SSiB (R), ECMWF (S), MOSES (T), and MOSES-CEH (U)), and one model in group 1 (RCA (B)). However, three of the models in group 4 underestimate the amount of snow during the melt period, while SSiB (R) simulates a greater peak accumulation than observed. The SPONSOR (A) simulations show a large melt event in early March in cells 1, 2, and 3, which is not matched in magnitude by any of the other models or the observations.

Most models predict a longer snow covered period in cell 4, the only cell without overstory vegetation, than in the other cells. Once again there is a large discrepancy among the models, both in the amount of accumulated snow and the occurrence of the first snow-free day. All models simulate the smallest snow accumulations for cell 5, but maximum accumulations range from about 150 mm for HY-SSiB (L) to less than 50 mm for a number of other models (e.g., RCA (B), MECMWF (O), NOAH (P), SSiB (R), and ECMWF (S)).

Fig. 4 shows the first simulated snow-free day for the same five cells used in Table 2 and Fig. 3 (note that the snow water equivalent measurements do not extend far enough into the snowmelt season to allow determination of the first snow-free day). As in Section 3.1, a lower threshold of 10-mm snow water equivalent is used. Most of the models simulate the first snow-free day for cells 1, 2, 3, and 5 to be during

the 3rd week of May, and the first snow-free day in cell 4 to be about 1 week later. However, there is considerable disagreement among the models, with a difference in the simulation of the first snow-free day of 8 weeks or more. Some models melt consistently earlier than the others (e.g., MATSIRO (K), HY-SSiB (L), NOAH (P), and ECMWF (S)), but other models only show a different behavior for one or two grid cells (e.g., SSiB (R) in cell 4).

Fig. 5 shows the partitioning of snowfall into melt and sublimation. Averaged over the entire basin and study period, snowfall formed 43% of the total precipitation (ranging from 34% in the southeast to 56% in the northwest). Total annual snowfall ranged from 212 mm in the southeast to 773 mm in the northwest, with a basin-wide average of 306 mm. The basin-wide, mean annual sublimation varied widely among the models, ranging from small negative amounts (condensation) for HY-SSiB (L) and SWAP (Q) to about 100 mm of sublimation for NOAH (P) and the models in Group 4. The melt in Fig. 5 is calculated as the difference between the prescribed snowfall and the simulated sublimation, assuming that the mean annual change in snow storage is negligible. MATSIRO (K) modified the specified snow and rain fractions, and the sum of melt and sublimation is consequently slightly less than 306 mm/year. Melt was calculated as a balance term, because the reported “melt”

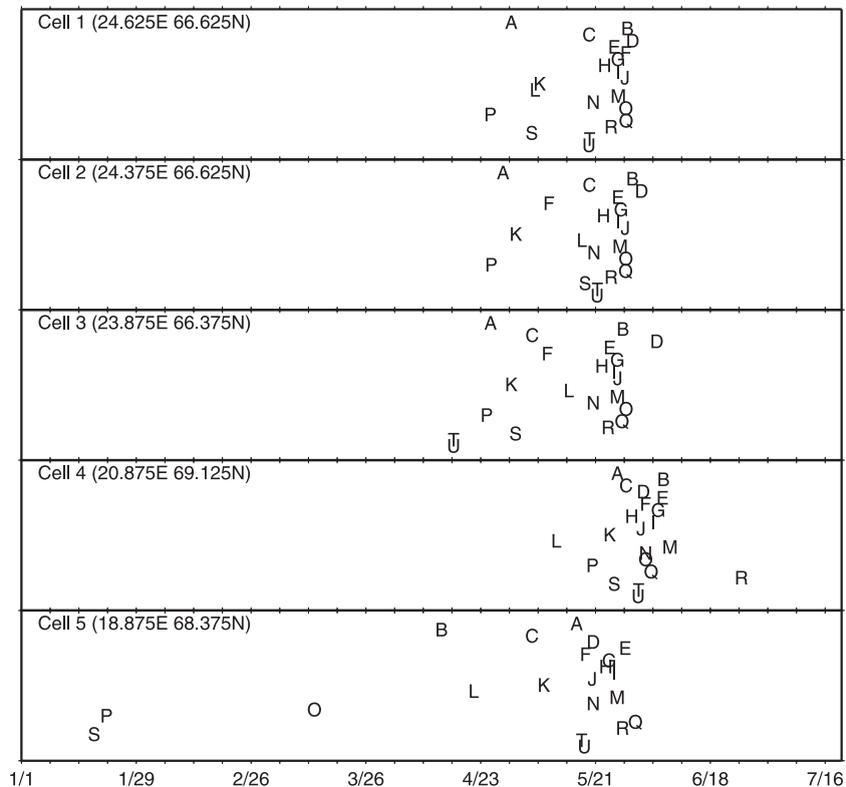


Fig. 4. First snow-free day in 1995 for five locations with different vegetation types.

variable was snowpack outflow, which included liquid precipitation that drained through the snowpack.

On a monthly timescale, most models simulate the highest snow sublimation during March, when mean areal sublimation rates vary from  $-0.03$  mm/day for HY-SSIB (L) to  $0.85$  mm/day for NOAH (P) (not shown). There are at least two possible reasons for the relatively high snow sublimations rates for NOAH (P). First, the use of the global  $1^\circ$  maximum snow albedo data set from Robinson and Kukla (1985) generally results in a lower albedo for the snow-covered area in the composite NOAH snow scheme compared to other models. Second, when snow is present in a NOAH grid cell, the snowpack is the only source of evaporation (sublimation), even if the grid box is only partially covered. As discussed in Section 8, the relatively high latent flux for ECMWF (S) is a result of the snow surface roughness parameterization. Reduction of the snow surface

roughness, mainly by allowing the snowpack to cover short vegetation leads to a reduction of the annual latent heat flux of about 25% or about 100 mm/year (van den Hurk and Viterbo, 2003-this issue).

#### 4. Streamflow

Daily discharge for the period 1979–1998 was provided for 20 stations throughout the Torne–Kalix basin by SMHI (Fig. 1). Fifteen of the stations were corrected by SMHI for the influence of ice jams for at least some of the years of record. Winter flows are estimated by SMHI, based on at least one flow meter measurement per winter season, using flow meters lowered through bored holes in the ice cover.

Both the Torne and the Kalix Rivers are essentially unregulated. A natural bifurcation on the Torne river diverts discharge above Junosuando/Tarendo to the

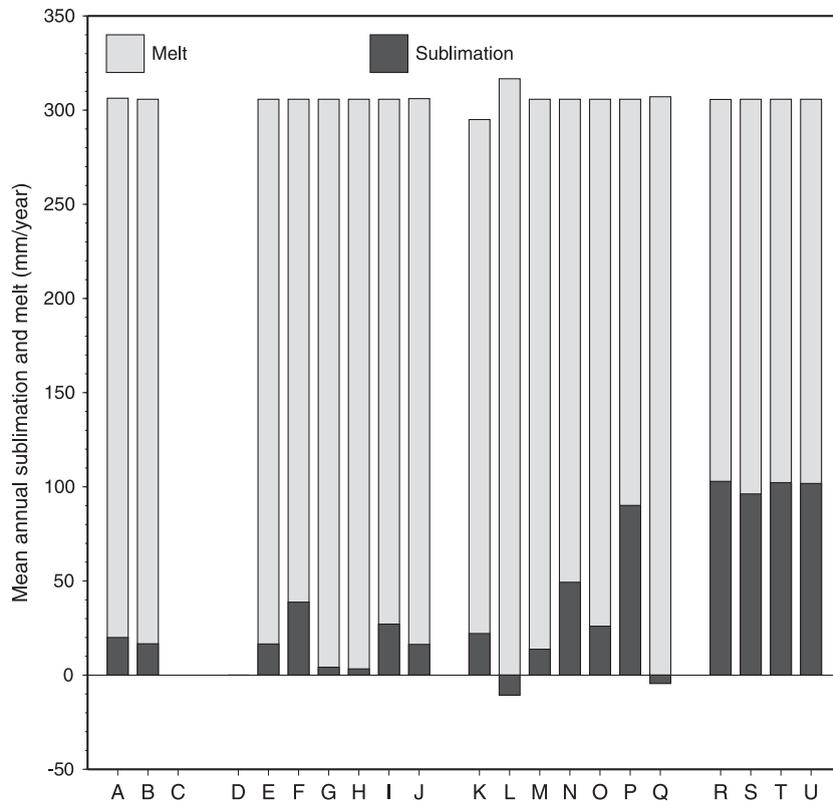


Fig. 5. Mean annual basin-wide sublimation and melt (both in mm). Note that IHAS (C) and SEWAB (D) did not provide snow sublimation.

Kalix River (see Fig. 2 in Bowling et al., 2003a-this issue). The bifurcation is gauged, so the effect of this diversion can be accounted for in the routing of simulated runoff. Observed flows at two gauges were used to determine the fraction of the flow diverted from the Torne to the Kalix river for each month during the study period. These monthly fractions were then used by the routing model to partition the daily simulated flows. The remaining 18 gauging stations were used for comparison with observations. Model-generated runoff at each of the individual grid cells was routed through a simplified river network using the same daily time step routing model used in the PILPS Phase 2(c) experiment (Lohmann et al., 1998). However, unlike in Phase 2(c), the grid cell runoff for each model was routed using the same routing model coefficients.

Fig. 6 shows the mean monthly simulated and observed discharge at the main gauging stations on the Torne and Kalix rivers. River discharge is char-

acterized by a strong snowmelt-related peak in May and June, followed by a gradual recession during the summer and fall. Throughout the winter, runoff is small, because snowmelt is negligible and most of the precipitation is stored in the snowpack. Mid-winter melt events are relatively uncommon at this latitude. Although the rivers are covered by ice during much of the winter, they continue to flow under the ice, although under-ice discharge is relatively small.

On a monthly time scale, all models capture the general pattern of a snowmelt peak in late spring/early summer, followed by a gradual recession during the remainder of the summer and the fall, and little or no runoff during the winter (Fig. 6). The among-model differences in the simulated runoff peak are closely related to the differences in snow accumulation and melt. The models that showed a shallow snowpack generally underestimate the streamflow, and some of the models with very deep snowpacks show very large runoff peaks. However, the snowmelt signal is modu-

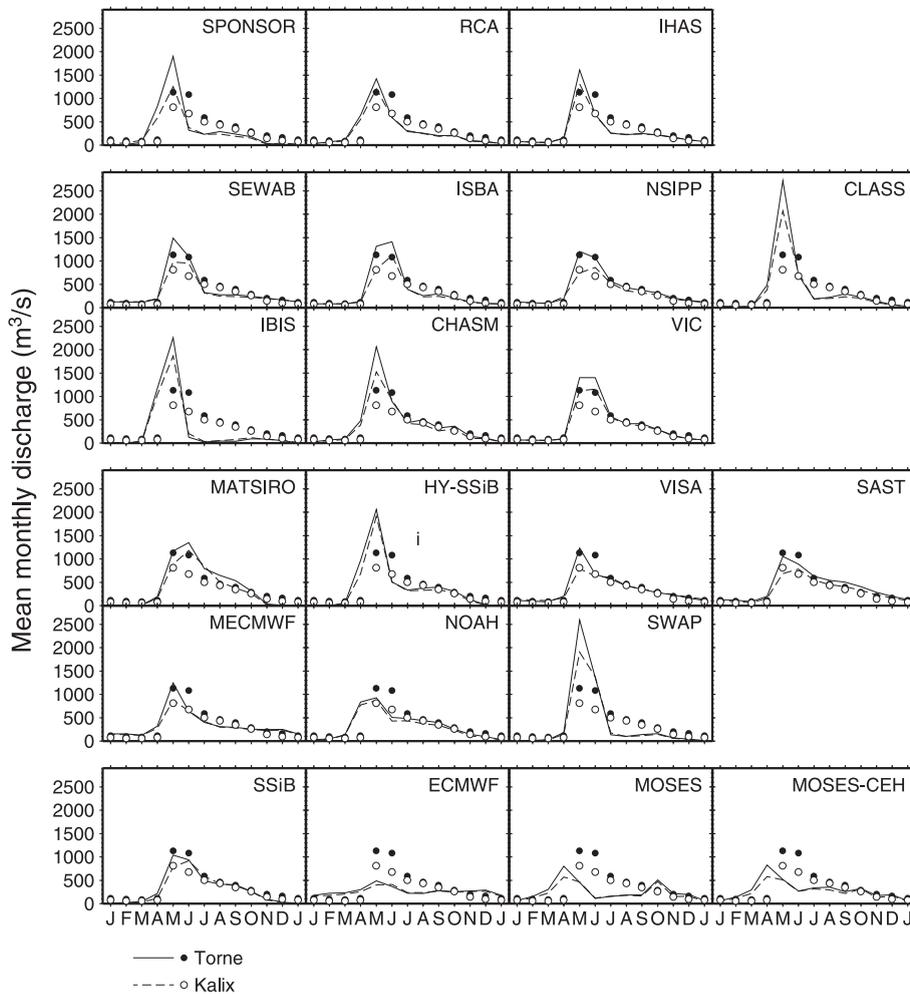


Fig. 6. Mean monthly observed (dots) and simulated (lines) discharge for the Kalix and Torne river basins.

lated by surface and subsurface storage, which is represented differently by the different models. For example, IHAS (C) and IBIS (H) have similar maximum snow accumulations (Fig. 3), but IBIS has a much larger discharge peak and less flow during the summer. Fig. 16 in Bowling et al. (2003a-this issue) shows the partitioning of mean annual runoff into surface and subsurface runoff for each model. The models that classify most of their runoff as surface runoff tend to have higher discharge peaks following snowmelt, and a steeper recession thereafter. Surprisingly, many of the models that showed more snow than observed in Fig. 3 match the observed streamflow

better than models whose maximum snow accumulations were close to the observed (notably ECMWF (S), MOSES (T), and MOSES-CEH (U)).

To further investigate the temporal and spatial differences in runoff production at the grid cell level, the monthly among-model standard deviation in grid cell runoff (calculated from the grid cell averages over the period of record for each model) is presented in Fig. 7a. Mean monthly runoff production is shown in Fig. 7b. During all months, the among-model standard deviation of runoff generation is of similar magnitude to the mean monthly runoff generation. The largest variability in runoff production occurs in May, when

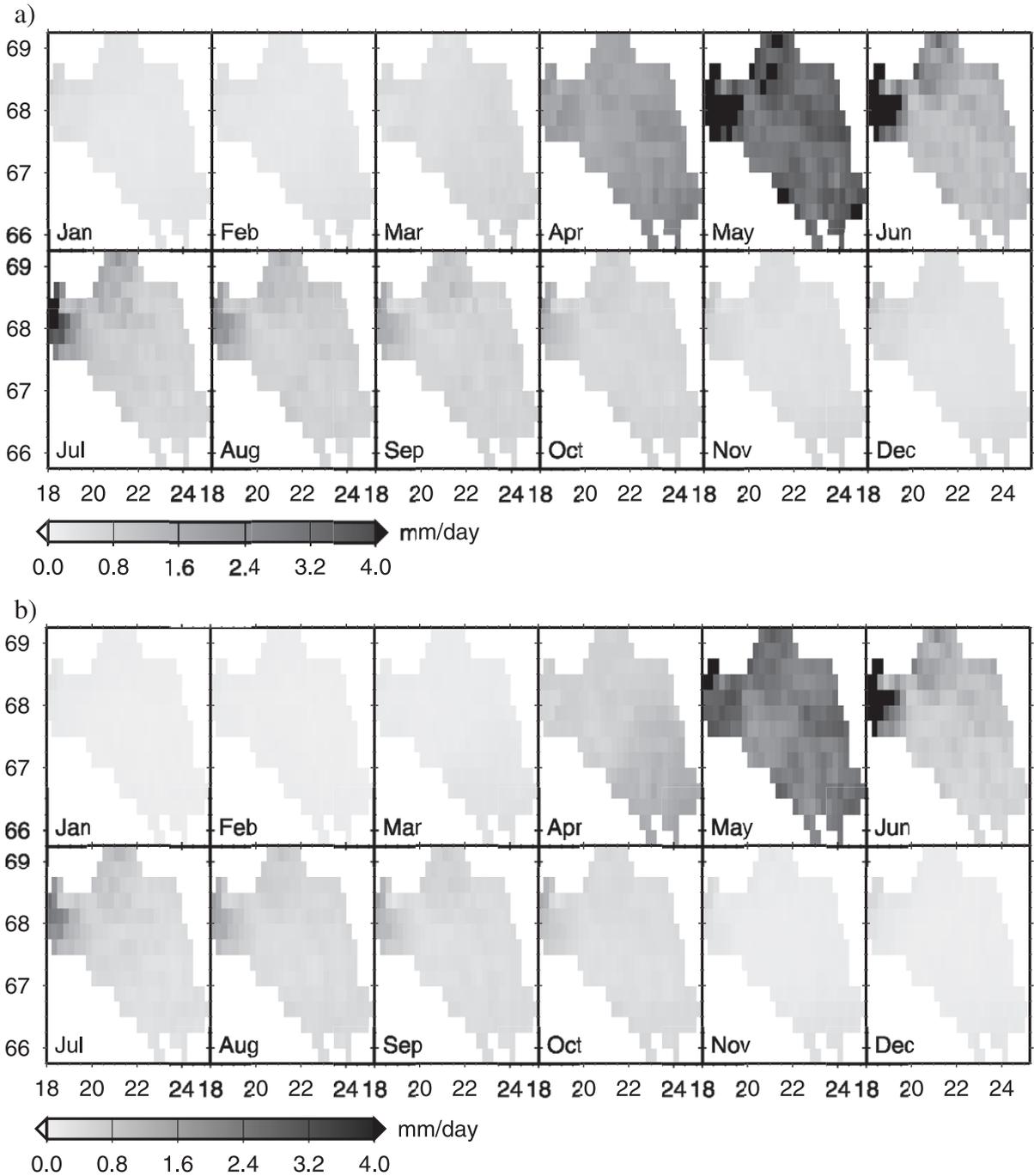


Fig. 7. (a) Between model variance in mean monthly runoff production at each grid cell. (b) Mean monthly runoff production (averaged over all models).

runoff production is largest. The large variability is a direct result of the differences in accumulated snow and the timing of the snowmelt. Outside of the snowmelt season mean runoff is less than 1 mm/day.

Fig. 8 shows the daily flow duration curves based on the entire study period for the observed and model simulated flows for the Kalix and Torne rivers. Observed and simulated daily flows were normalized by the upstream area to reduce errors due to discrepancies between the true and model drainage networks.

Flow duration curves show the frequency with which discharge thresholds are exceeded and as such are the complement of cumulative distribution functions. Note that the lowest observed flows occur during the winter when snowmelt is negligible. Many models show higher peak flows than observed and smaller winter flows than observed, which corresponds to a higher variance of the simulated daily streamflow compared to the observed. This is also demonstrated in Fig. 9, which shows a Taylor diagram of the daily simulated

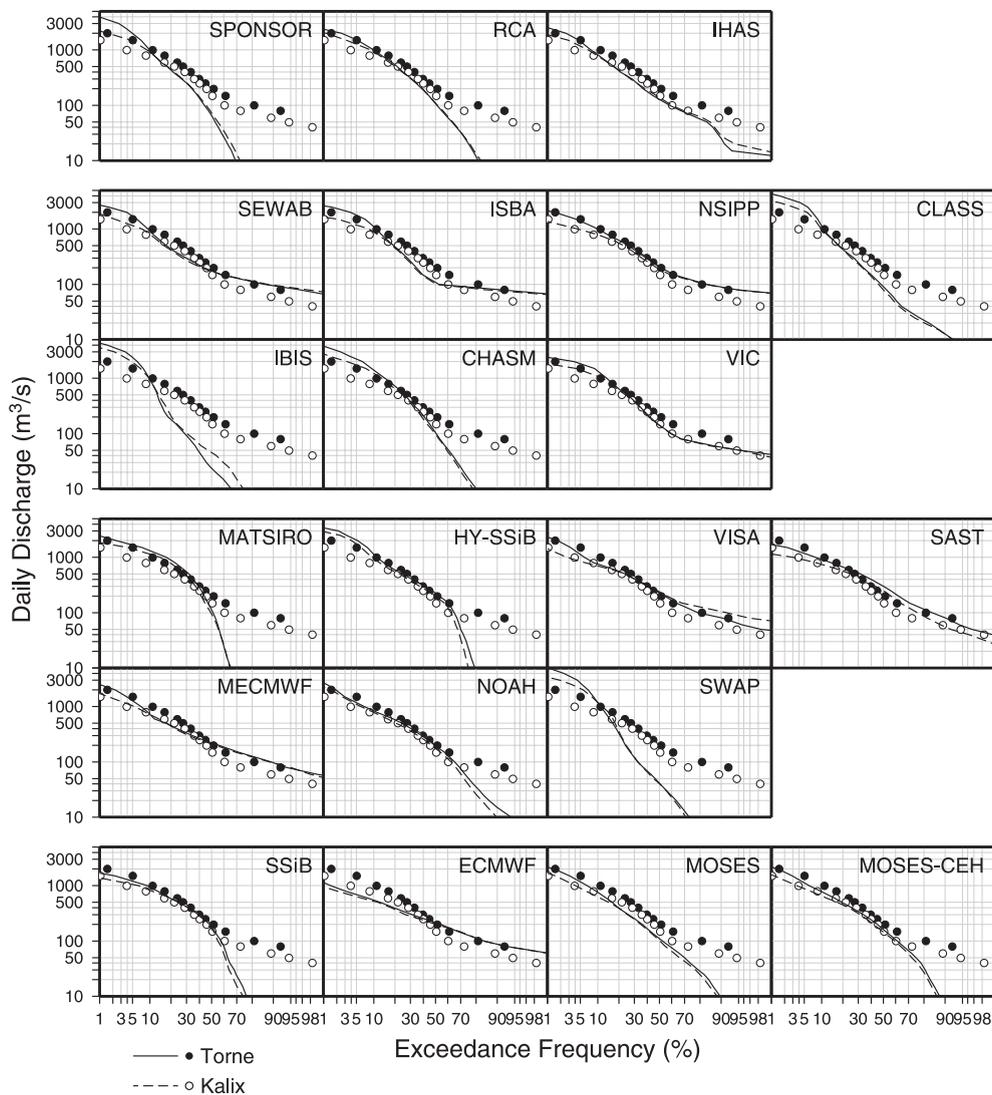


Fig. 8. Flow duration curves based on observed (dots) and simulated (lines) daily flows during the period 1989–1998.

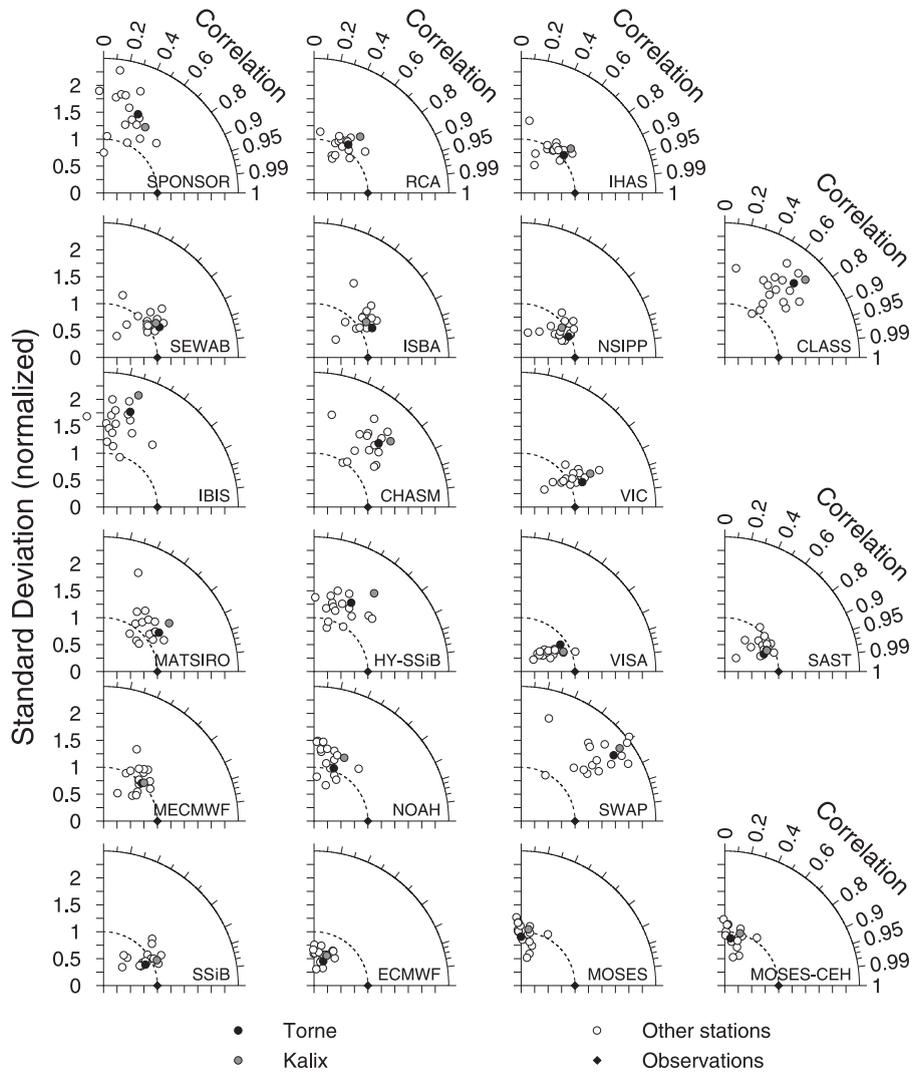


Fig. 9. Taylor diagram of daily flows during the period 1989–1998. See text for details.

flows for each model. The Taylor diagram allows the representation of three important second-order statistics of the simulated series to be reflected in one single plot (McAvaney et al., 2001). The distance from the origin to a plotted point is the ratio of the standard deviations of the simulated and observed flows, while the cosine of the azimuth angle is the correlation between the observed and simulated daily flows, after removal of the mean. The observed flows consequently plot along the horizontal axis, at unit distance from the origin. The third statistic is that part of the

root mean squared error (RMSE) that is not attributable to bias in the simulated flows. This is given by the distance from the point corresponding to the observed flow to any other plotted point. The Taylor diagram reflects only the random errors in the daily flow calculations. However, even if the random error tends to zero, a large systematic error or bias may still exist in the simulated flows. Table 3 shows in summary form the systematic components of the errors in the simulated flow (bias and root mean squared error (RMSE)).

Table 3

Summary statistics of daily simulated flows compared to observed flows for 18 stations in the Torne and Kalix river basins (all values in mm/day)

Model	All river gauges						Torne		Kalix	
	Bias			RMS Error			Bias	RMSE	Bias	RMSE
	Min	Median	Max	Min	Median	Max				
SPONSOR	-0.658	-0.200	-0.062	1.445	2.838	5.143	-0.190	1.844	-0.131	1.526
RCA	-0.643	-0.273	0.026	1.036	1.671	3.428	-0.220	1.199	-0.009	1.290
IHAS	-0.709	-0.218	-0.005	0.926	1.423	3.767	-0.256	0.926	-0.039	1.018
SEWAB	-0.396	-0.070	0.095	0.667	1.107	3.129	-0.086	0.698	0.063	0.781
ISBA	-0.571	-0.132	0.007	0.686	1.118	3.163	-0.121	0.686	-0.033	0.807
NSIPP	-0.554	-0.044	0.123	0.490	0.924	2.615	-0.049	0.499	0.084	0.749
CLASS	-0.527	-0.007	0.211	1.375	2.229	3.947	0.024	1.715	0.184	1.871
IBIS	-0.590	-0.211	0.017	1.953	2.981	5.145	-0.226	2.244	-0.029	2.573
CHASM	-0.424	0.101	0.319	1.008	2.044	3.941	0.105	1.460	0.288	1.607
VIC	-0.394	0.005	0.272	0.584	0.981	2.104	0.001	0.584	0.229	0.860
MATSIRO	-0.483	-0.052	0.231	0.814	1.567	2.502	0.003	0.881	0.189	1.146
HY-SSiB	-0.442	0.044	0.343	1.415	2.157	4.343	0.031	1.599	0.271	1.808
VISA	-0.438	-0.116	0.150	0.537	0.991	2.233	-0.131	0.709	0.121	0.537
SAST	-0.442	0.014	0.201	0.517	0.882	2.846	-0.004	0.517	0.106	0.560
MECMWF	-0.567	-0.111	0.095	0.880	1.612	2.774	-0.132	0.937	0.075	0.930
NOAH	-0.672	-0.169	0.089	1.323	2.414	4.863	-0.158	1.432	0.053	1.536
SWAP	-0.936	-0.044	0.229	1.362	2.294	4.140	-0.033	1.729	0.137	1.945
SSiB	-0.700	-0.247	0.076	0.584	0.930	2.905	-0.230	0.597	-0.013	0.584
ECMWF	-1.001	-0.436	-0.105	1.164	1.789	3.350	-0.368	1.200	-0.174	1.164
MOSES	-0.957	-0.388	-0.255	1.476	2.308	4.340	-0.369	1.684	-0.255	1.676
MOSES-CEH	-0.884	-0.314	-0.148	1.308	2.179	4.142	-0.278	1.548	-0.157	1.483

The models that have flow duration curves similar to the observed, plot near the dashed arc, which means that their daily flows have a similar standard deviation to the observed flows. The ECMWF (S) model plots below this arc, because it shows much less variability than the observed flows, with higher flows than observed during the winter, and lower peak flows in spring. Because the main runoff event is snowmelt related, the models that ablate their snowpacks early (Fig. 2), are poorly correlated with the observed flows on a daily time step (NOAH (P), ECMWF (S), MOSES (T), and MOSES-CEH (U)), and plot in the left part of the diagram. Models that have very high snowmelt related peaks (Fig. 6) have a greater variance than the observed flows (SPONSOR (A), CLASS (G), IBIS (H), CHASM (I), HY-SSiB (L), and SWAP (Q)), and plot along the outer arc. Other than HY-SSiB (L), all these models generate most of their annual runoff as surface runoff.

As mentioned in Bowling et al. (2003a-this issue), those models that participated in the calibration experiment and subsequently transferred their cali-

bration results from the subbasins to the region at large, generally showed better daily flow simulations. Of the 9 models in Fig. 8 with good low flow simulations (IHAS, SEWAB, ISBA, NSIPP, VIC, VISA, SAST, MECMWF, ECMWF), only 2 (IHAS, ECMWF) did not use transferred calibration results. Put differently, of the 10 models that used transferred calibration results, 7 produced good low flow simulations. Of the remaining 3 (MATSIRO, HY-SSiB, SSiB), at least 1 (HY-SSiB) underpredicted low flows because no baseflow was produced when the soil was frozen. Nine out of 11 models that did not use calibration results underpredicted winter low flows (Fig. 8). As a group, the 10 models that transferred calibration results had a mean absolute bias and mean RMSE of 0.079 and 0.771 mm/day, respectively, for the Torne river and 0.118 and 0.876 mm/day for the Kalix river. In contrast, the 11 models that did not use calibration results had a mean absolute bias and mean RMSE of 0.202 and 1.544 mm/day, respectively, for the Torne river and 0.132 and 1.608 mm/day for the Kalix river.

## 5. Surface energy balance

As discussed in [Bowling et al. \(2003a-this issue\)](#), all models except IHAS (C) close their mean annual energy balance to within  $3 \text{ W/m}^2$ . It is somewhat surprising that RCA (B) closes its energy balance, despite the discrepancy in the water balance (Section 2), since the energy and water balance are tightly linked through the latent heat flux.

Estimated basin-wide runoff from observations is about  $403 \text{ mm/year}$  during the study period, while basin-average observed and gauge-catch corrected precipitation during the same period is about  $704 \text{ mm/year}$  ([Bowling et al., 2003a-this issue](#)). Assuming that the cumulative basin-wide storage change during the study period is small compared to the other terms in the water balance, basin average evaporation plus sublimation is

about  $300 \text{ mm/year}$ . This corresponds to a mean annual latent heat flux between  $24$  and  $27 \text{ W/m}^2$ , depending on the ratio of sublimation to evaporation.

The sum of simulated evaporation and sublimation ranges from  $221 \text{ mm/year}$  for IBIS (H) to  $418 \text{ mm/year}$  for MOSES (T). Seven models have totals smaller than  $270 \text{ mm/year}$ , six models have totals greater than  $330 \text{ mm/year}$ , and eight fall in between. Three of the models in group 4 (ECMWF (S), MOSES (T), and MOSES-CEH (U)) have relatively large sublimation during the winter and spring. NOAH (P) and SSiB (R), which both have relatively high sublimation rates during the spring, have smaller evaporation maxima in the summer than most of the other models. Mean monthly evaporation reaches a maximum for all models in June or July, but maximum evaporation rates range from  $1.7 \text{ mm/day}$  for SAST (N) to  $3.2 \text{ mm/day}$  for SWAP (Q)

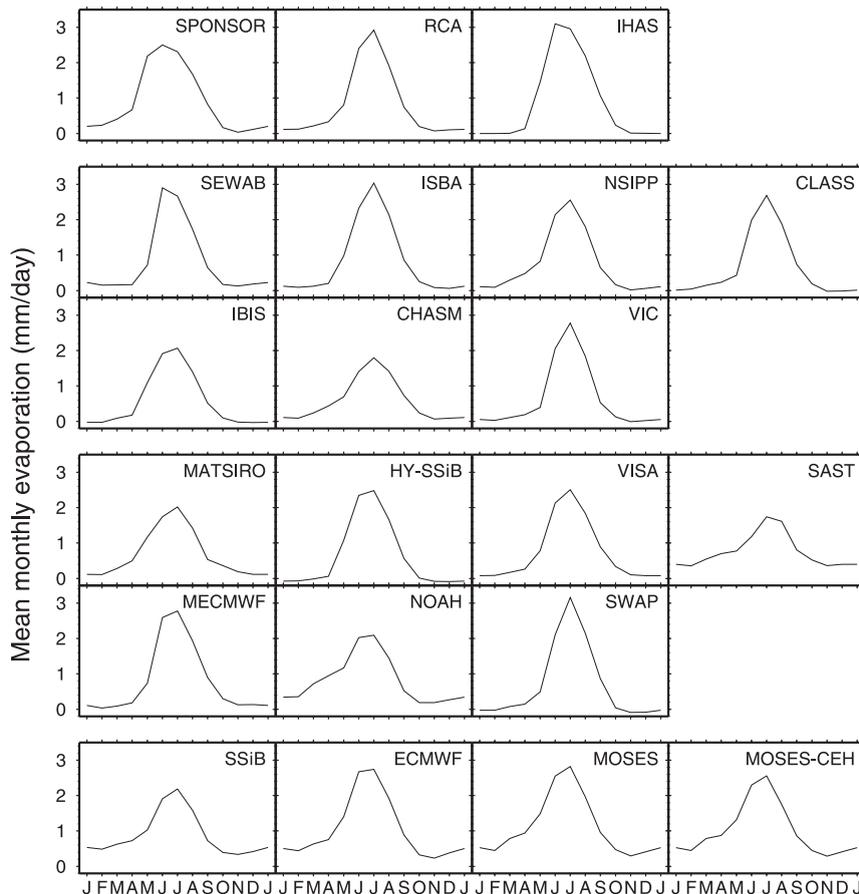


Fig. 10. Mean monthly simulated evaporative flux from all sources, including sublimation.

(Fig. 10). Note that SAST (N) and SWAP (Q) have similar amounts of annual runoff, but that SWAP (Q) shows a much more peaked runoff response (see Figs. 6, 8 and 9).

Fig. 11 shows the mean annual basin-wide net radiation fluxes for each of the models. The mean annual downward shortwave radiation and downwelling longwave radiation were 90.2 and 268 W/m<sup>2</sup>, respectively. Because most of the basin lies above the Arctic Circle, downward shortwave radiation is very small during the winter months. Mean monthly basin-wide downward shortwave radiation is less than 1 W/m<sup>2</sup> in December and January, but is more than 200 W/m<sup>2</sup> in June and July. Downwelling longwave radiation varied from 223 W/m<sup>2</sup> in February to 331 W/m<sup>2</sup> in July. The mean annual simulated net radiation is generally small and varied between 16.4 W/m<sup>2</sup> (SWAP (Q)) and 40.4 W/m<sup>2</sup> (IHAS (C)), with a mean of 26.4 W/m<sup>2</sup>. The mean monthly net radiation (not shown) peaked for all models in June, with mean

monthly net radiation for June between 106 W/m<sup>2</sup> (SWAP (Q)) and 156 W/m<sup>2</sup> (IHAS (C)). For all models, net radiation is negative from October through February. Net shortwave radiation shows the largest variability, both among models and among grid cells for a single model, during April and May. This is the period when incoming shortwave radiation is increasing and differences in surface albedo are large as a result of differences in both snow (and snow-free) albedo and snow cover.

The ground heat flux and phase change fluxes are small averaged over a year, and consequently the sum of the mean annual sensible and latent heat fluxes is approximately equal to the mean annual net radiation. Because the all-model mean annual net radiation of 26.4 W/m<sup>2</sup> is of similar magnitude to the all-model mean annual latent heat flux (model average 25.0 W/m<sup>2</sup>), the mean annual sensible heat flux is small, varying from -10.5 to +8.5 W/m<sup>2</sup>, with a model average of -1.0 W/m<sup>2</sup> (Fig. 12). This

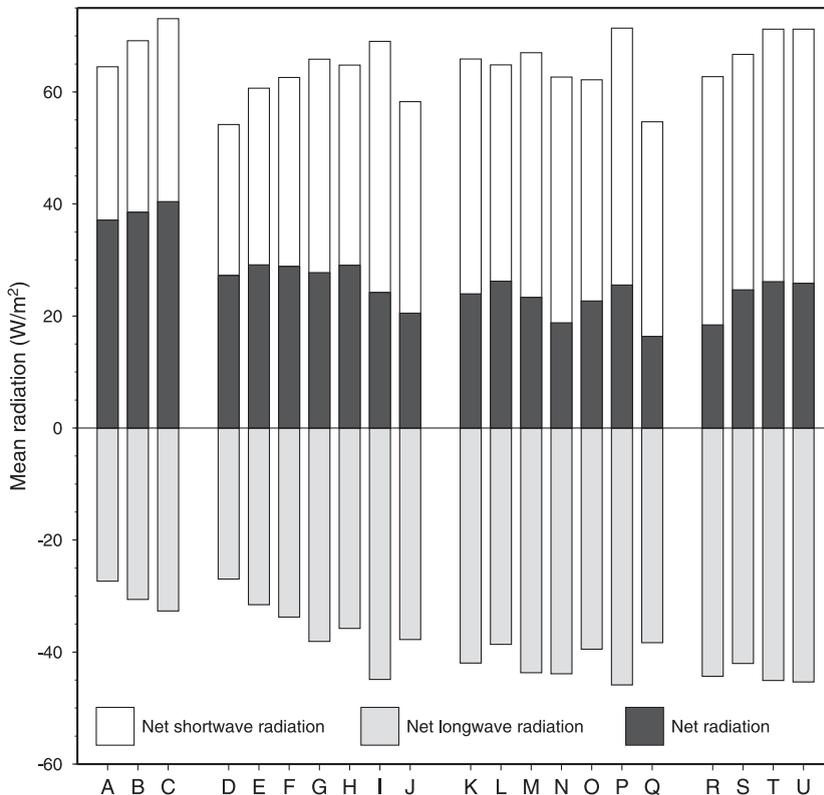


Fig. 11. Mean annual basin-wide net shortwave radiation, net longwave radiation and net radiation.

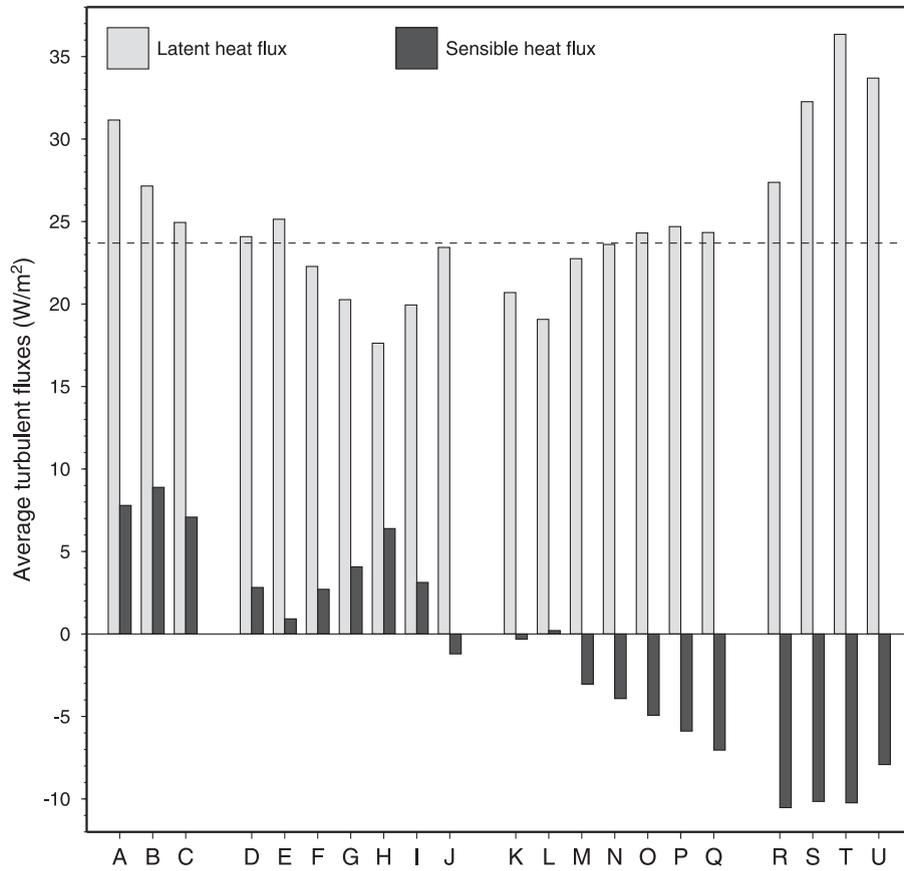


Fig. 12. Mean annual basin-wide turbulent heat fluxes. The dashed line shows the water balance based estimate of the latent heat flux over the basin (see Bowling et al., this issue a).

small annual sensible heat flux is the net result of a positive (away from the land surface) sensible heat flux in spring and summer and a negative sensible heat flux during the fall and winter. Generally, models that have large amounts of sublimation (Fig. 5) have a negative sensible heat flux, suggesting that for these models the atmosphere acts as a source of energy for the winter-time sublimation.

## 6. Surface temperature and albedo

Downward shortwave and downwelling longwave radiation were provided as model forcings and the thermal emissivity was prescribed as 1.0 for all

surfaces. Consequently, between-model differences in net radiation are a result of differences in shortwave albedo and surface radiative temperature.

Net radiation ( $R_{\text{net}}$ ) is defined as

$$\begin{aligned} R_{\text{net}} &= (R_{\text{S}}^{\downarrow} - R_{\text{S}}^{\uparrow}) + (R_{\text{L}}^{\downarrow} - R_{\text{L}}^{\uparrow}) \\ &= (1 - \alpha)R_{\text{S}}^{\downarrow} + (R_{\text{L}}^{\downarrow} - \varepsilon\sigma T_{\text{R}}^4), \end{aligned} \quad (1)$$

where  $R_{\text{S}}^{\downarrow}$  and  $R_{\text{S}}^{\uparrow}$  are the downward and reflected shortwave radiation, respectively,  $R_{\text{L}}^{\downarrow}$  and  $R_{\text{L}}^{\uparrow}$  are the downwelling and emitted longwave radiation,  $T_{\text{R}}$  is the surface radiative temperature,  $\alpha$  is the shortwave surface albedo,  $\varepsilon$  is the surface emissivity, and  $\sigma$  is the Stefan–Boltzmann constant.

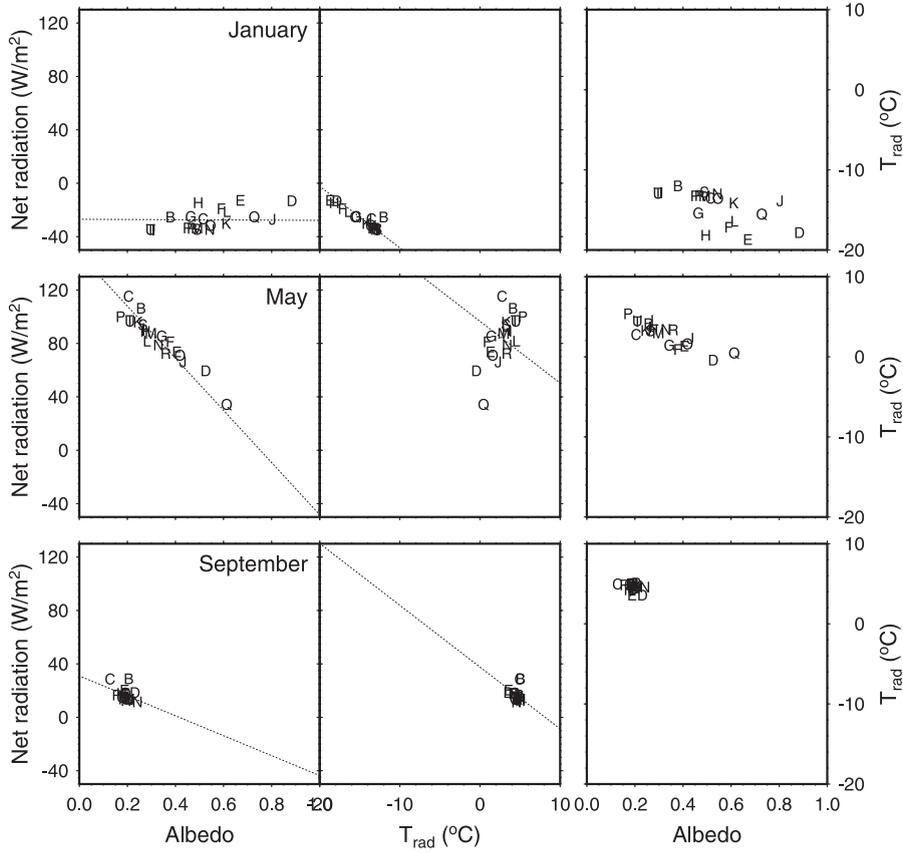


Fig. 13. Mean monthly net radiation versus albedo and radiative surface temperature and surface temperature versus albedo for January, May, and September. The slope of the dashed lines in the first two columns shows how net radiation would change if either albedo or radiative temperature would be the only control on net radiation. See text for details.

The sensitivity of net radiation to changes in albedo is given by

$$\frac{\partial R_{\text{net}}}{\partial \alpha} = -R_{\text{S}}^{\downarrow} \quad (2)$$

Thus, at high northern latitudes the effect of albedo on net radiation is small during the winter months when incoming radiation is close to 0 W/m<sup>2</sup>, and is large from late spring through early fall. Similarly, we can determine the effect of the radiative temperature on net radiation as

$$\begin{aligned} \frac{\partial R_{\text{net}}}{\partial T_{\text{R}}} &= -4\epsilon\sigma T_{\text{R}}^3 = -4\epsilon\sigma(T_{\text{R}}^{\text{c}} + 273.15)^3 \\ &\approx -4\epsilon\sigma 273.15^3, \end{aligned} \quad (3)$$

where  $T_{\text{R}}^{\text{c}}$  is the radiative temperature in °C, which in most cases is small compared to 273.15. Because the

emissivity of all surfaces was set to 1, the sensitivity of net radiation to the surface radiative temperature is about  $-4.6 \text{ W/m}^2/\text{°C}$ .

These simple relationships can be used to investigate the effects of differences in albedo and surface radiative temperature on model simulated net radiation. In particular, they can be used to determine whether the differences in simulated net radiation are mainly due to differences in albedo or to differences in simulated radiative temperature.

Fig. 13 shows mean monthly net radiation versus albedo and radiative surface temperature for January, May and September. The dashed lines in the left and middle columns have slopes given in Eqs. (2) and (3), respectively. Thus, the slopes in the left column are  $-0.8$ ,  $-195$ , and  $-75 \text{ W/m}^2$ , from top to bottom. All the slopes in the middle column are  $-4.6 \text{ W/m}^2/$

°C. To illustrate the interdependence of albedo and surface temperature, a plot of surface temperature versus albedo is shown in the right column.

The sensitivity of net radiation to surface albedo is low when incoming shortwave radiation is low. Consequently, most of the among-model differences in net radiation in January are attributable to differences in surface radiative temperature. This is demonstrated in the middle panel, where all models except two plot along a line with a slope of  $-4.6 \text{ W/m}^2$ . Decreasing radiative temperatures result in higher net radiation. The deviation from the line for RCA (B) results from their use of a different snow emissivity (0.95) than prescribed (1.0). Because of the insensitivity during January of net radiation to shortwave albedo, there is essentially no relationship between surface temperature and albedo. Interestingly, the four models with the lowest surface albedo in January (RCA (B), CHASM (I), MOSES (T), and MOSES-CEH (U)) are the only four models that have a composite snow model, that is, the snow layer is incorporated into the soil model.

The situation is very different in May, when incoming shortwave radiation is high ( $195 \text{ W/m}^2$ ) and small changes in shortwave albedo result in large changes in net radiation. For example, a change in albedo of 0.05 has the same effect as a change in radiative temperature of about  $2 \text{ }^\circ\text{C}$ . Mean monthly radiative surface temperatures range from  $-0.46$  to  $+5.33 \text{ }^\circ\text{C}$  in May. This range in temperatures could explain a variation in net radiation of about  $27 \text{ W/m}^2$ . However, the range in simulated shortwave albedo values (0.171–0.613) is equivalent to a much larger

change in net radiation of about  $86 \text{ W/m}^2$ . Albedo and radiative temperature show a much stronger relationship in May than in January. Higher albedos are coincident with lower radiative temperatures because they indicate the presence of snow. Because of the large sensitivity of net radiation to albedo and the relationship between albedo and radiative temperatures, higher surface temperatures are coincident with higher net radiation as shown in the middle panel. In May, among-model variation in net radiation is driven by differences in snow cover and hence albedo.

In September, the situation is different once again. Snow does not play a role in the land-surface energy balance. Albedos, radiative temperatures and net radiation vary relatively little among the models. Both albedo and radiative temperature are negatively correlated with net radiation. Albedos range from 0.128 to 0.242 across the models, equivalent to a difference in net radiation of  $8.5 \text{ W/m}^2$ , and temperatures range from  $3.34$  to  $5.06 \text{ }^\circ\text{C}$ , equivalent to a difference in net radiation of about  $8 \text{ W/m}^2$ . Average simulated net radiation was  $17.4 \text{ W/m}^2$  with a range from  $11.3$  to  $28.7 \text{ W/m}^2$ .

## 7. Surface temperature and the sensible heat flux

Surface temperature affects not only the emitted longwave radiation, but is also an important factor in the sensible heat flux. Sensible heat is directly dependent on the temperature gradient near the surface, and increases with the difference between the land-surface temperature and the air temperature.

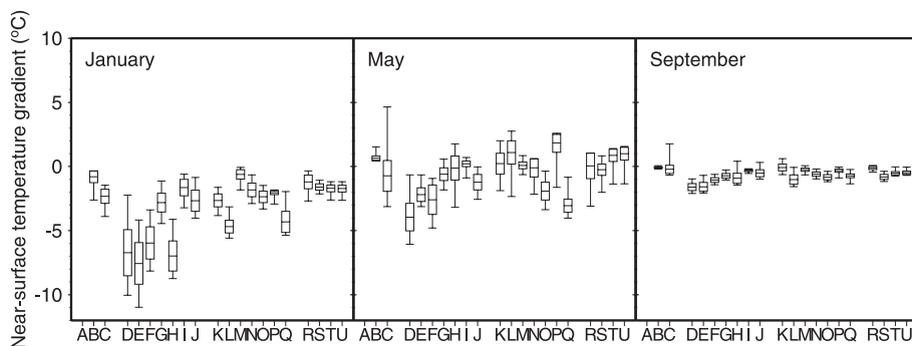


Fig. 14. Surface temperature depression (land-surface temperature minus air temperature) for all models for January, May, and September. Each bar and whisker shows the mean, 1 S.D. on both sides of the mean, and the minimum and maximum value for all 218 model grid cells.

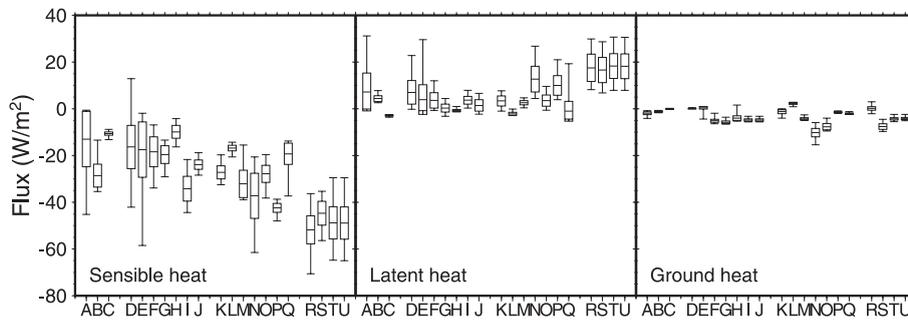


Fig. 15. Box and whisker plots for the sensible, latent, and ground heat flux in January. Each bar and whisker shows the mean, 1 S.D. on both sides of the mean, and the minimum and maximum value for all 218 model grid cells.

Most models use a stability correction that increases turbulent exchange under unstable conditions (land surface warmer than the air temperature) and limits the sensible heat flux under extremely stable conditions (land surface colder than the air temperature). Slater et al. (2001), based on results from the PILPS Phase 2(d) model intercomparison at Valdai, Russia, argued that in some models the land surface becomes decoupled from the atmosphere during the coldest and darkest months as a result of this mechanism. Basically, the hypothesis is that low surface temperatures lead to extremely stable conditions in which the turbulent fluxes, and in particular the sensible heat flux, tends to zero. Because net radiation is negative during the winter, the land-surface temperature would drop further through radiative cooling. This feedback mechanism leads to a very large difference between the surface temperature and the air temperature during the winter as compared to during the summer. Fig. 14 shows the surface temperature depression (surface temperature minus air temperature) for all models during January, May, and September. The box and whisker plots show the mean surface temperature depression, 1 S.D. each side of the mean, and the minimum and maximum values for all of the 218 grid cells.

Most models show a temperature gradient between 0 and  $-5$  °C (land surface colder than the air) during January. However, four models (SEWAB (D), ISBA (E), NSIPP (F), and IBIS (H)) are significantly colder. Fig. 15 shows the variation in the sensible, latent, and ground heat flux during January. The four models with the largest temperature depression in January do not stand out as a group with respect to turbulent and

ground heat fluxes. Because the sensible heat flux is a function not only of the near surface temperature gradient, but also of the surface roughness and the parameterization of the stability correction, similar temperature gradients can result in different sensible heat fluxes. In addition, the monthly mean grid cell values used to construct Figs. 14 and 15 may mask brief periods of unstable conditions, when most of the turbulent flux exchange may occur. The strongest pattern in Fig. 14 is that the magnitude of the near surface temperature gradient becomes smaller as the snow disappears. Similarly, the variability in the near surface temperature gradient, both the among-grid cell variability and the among-model variability, diminishes as the snow disappears. During January, there is a large variability in the sensible heat flux, both among grid cells and among models.

The models that have a large positive latent heat flux (sublimation) have a large negative sensible heat flux. The magnitude of the ground heat flux is generally small compared to the turbulent fluxes, even for the models with the lowest land-surface temperatures. This is different from the findings of Slater et al. (2001), who found that the decoupling of the land surface and the atmosphere led to a large ground flux that matched the size of the net radiation. Although this may be the case for the coldest grid cell in the coldest models, this does not seem to be a general feature.

## 8. Surface roughness and turbulent fluxes

Most models used vegetation-dependent roughness lengths as recommended by the PILPS Phase 2(e)

experiment plan, combined with some form of stability correction (IHAS (C) was the only model that did not apply a stability correction). No roughness length was defined for snow and models differed in the way that snow-vegetation interactions were represented. The significance of differences in prescribed roughness is illustrated by the differences in simulated model fluxes between ECMWF (S) and MECMWF (O). These are essentially the same models, except that in MECMWF (O) some changes have been made to the runoff parameterization, the soil hydraulic properties, the snow surface roughness and the snow surface albedo (van den Hurk and Viterbo, 2003-this issue). In the modified model (MECMWF (O)), the surface roughness is a function of snow depth for short vegetation. Short vegetation can be completely covered by deep snowpacks, reducing the roughness length and the high wintertime latent heat flux simulated by ECMWF (S). The reduction in snow roughness leads to a reduction in the annual latent heat flux of about 25% (100 mm/year), partly because of a reduction in wintertime snow sublimation and partly because of reduced spring evaporation. The different behaviors of the two models are further explored by van den Hurk and Viterbo (2003-this issue). Similarly, the RCA (B) model had an unrealistically large wintertime snow sublimation in its original submission, which was greatly reduced in the final submission through a reduction of the snow surface roughness (Samuelsson et al., 2003-this issue). The large response of the wintertime snow sublimation to the snow surface roughness, and the consequent effects on the surface energy and water balance, demonstrates how relatively small changes in land-surface parameters can have more significant effects on simulated fluxes than differences in model structure.

## 9. Conclusions

All models participating in PILPS Phase 2(e) capture the broad dynamics of snowmelt and runoff, but large differences in snow accumulation and ablation, turbulent heat fluxes, and streamflow exist. One of the difficulties in interpreting the results from the PILPS Phase 2(e) experiments is the complexity of the current generation of land-surface schemes. Even in an experiment where meteorological forcings and many of the land-surface characteristics were prescribed, the

remaining number of degrees of freedom is large. Because of the nonlinearity of many of the land-surface processes, small differences in model parameters and in model parameterizations can lead to large differences in model outcomes (e.g., Takayabu et al., 2001). For example, differences in the parameterizations of grid cell fractional snow coverage result in differences in grid cell albedo and consequently in net radiation, even if the land surface and snow albedos are the same among the models. Differences in net radiation lead to differences in melt and turbulent exchange. Different parameterizations of land-surface roughness can result in large differences in the latent and sensible heat fluxes, and lead to substantial changes in both the water and energy balance terms as illustrated clearly by the differences between the simulations of the MECMWF (O) and ECMWF (S) models. In an interesting series of experiments using the CHASM (I) model, which can be operated with different complexities of the land-surface energy balance, Xia et al. (2003-this issue) demonstrate that the complexity of the representation of the land-surface energy balance does not explain the differences between the PILPS Phase 2(e) model results.

Nonetheless, some general conclusions can be drawn from the PILPS Phase 2(e) experiment.

- The greatest among-model differences in energy and moisture fluxes in these high-latitude environments occur during the spring snowmelt period, and reflect different model parameterizations of snow processes. Differences in net radiation are governed by differences in the radiative surface temperature during the winter months and by differences in surface albedo during the spring/early summer. Differences in net radiation are smallest during the late summer when snow is absent. Therefore, differences in annual net radiation (and hence other terms in the energy balance) are dominated by differences in the simulation of wintertime processes.
- Differences in simulated snow sublimation are an important source of among-model differences in water and energy balances. Even though sublimation is small for most models, a few models produce annual snow sublimation of about 100 mm. These differences in sublimation are reflected in other terms in the energy balance. For instance, models

with high sublimation also tend to have large negative sensible heat fluxes, both averaged over the winter period and averaged over the year. Furthermore, models with high sublimation tend to have below average spring snow accumulation, and mostly tend to underestimate spring and annual runoff. The differences in snow sublimation among models appear to be largely dependent on differences in snow surface roughness parameterizations.

- Differences in runoff parameterizations are largely reflected in differences in daily runoff statistics. Most models simulate runoff that is more variable than the observations at the daily time step. The models with the greatest variability, as much as twice the observed variability in daily flows, produce most of their runoff through fast response, surface runoff mechanisms. Although direct observations are lacking, it appears that much of the observed runoff is produced via subsurface or other “slow” response mechanisms (e.g., detention of runoff in lakes and wetlands), which are not well represented by many of the models.
- As a group, those models that participated in a calibration experiment in which calibration results were transferred from small catchments to the basin at large had a smaller bias and RMSE in their daily streamflow simulations than the group of models that did not.

## References

- Bowling, L.C., Lettenmaier, D.P., Nijssen, B., Graham, L.P., Clark, D.B., El Maayar, M., Essery, R., Goers, S., Habets, F., van den Hurk, B., Jin, J., Kahan, D., Lohmann, D., Mahanama, S., Mocko, D., Nasonova, O., Samuelsson, P., Shmakin, A.B., Takata, K., Verseghy, D., Viterbo, P., Xia, Y., Ma, X., Xue, Y., Yang, Z.-L., 2003a. Simulation of high latitude hydrological processes in the Torne–Kalix basin: PILPS Phase 2(c): 1. Experiment description and summary intercomparisons. *Glob. Planet. Change* 38, 1–30.
- Bowling, L.C., Lettenmaier, D.P., Nijssen, B., Polcher, J., Koster, R.D., Lohmann, D., 2003b. Simulation of high-latitude hydrological processes in the Torne–Kalix basin: PILPS Phase 2(c): 3. Equivalent model representation and sensitivity experiments. *Glob. Planet. Change* 38, 55–71.
- Dahlstrom, B., 1995. Snow cover. In: Raab, B., Vedin, H. (Eds.), *Climate, Lakes and Rivers, The National Atlas of Sweden*. SNA Publishing, Stockholm, pp. 91–111.
- Henderson-Sellers, A., Pitman, A.J., Love, P.K., Irannejad, P., Chen, T., 1995. The Project for Intercomparison of Land surface Parameterization Schemes (PILPS) Phases 2 and 3. *Bull. Am. Meteorol. Soc.* 76, 489–503.
- Lohmann, D., Lettenmaier, D.P., Liang, X., Wood, E.F., Boone, A.A., Chang, S., Chen, F., Dai, Y., Desborough, C., Dickinson, R.E., Duan, Q., Ek, M., Gusev, Y.M., Habets, F., Irannejad, P., Koster, R., Mitchel, K.E., Nasonova, O.N., Noilhan, J., Schaake, J., Schlosser, A., Shao, Y., Shmakin, A.B., Verseghy, D., Warrach, K., Wetzel, P., Xue, Y., Yang, Z.-L., Zeng, Q., 1998. The Project for Intercomparison of Land-surface Parameterization Schemes (PILPS) Phase 2(c) Red-Arkansas River Basin experiment: 3. Spatial and temporal analysis of water fluxes. *Glob. Planet. Change* 20, 161–179.
- McAvaney, B., Covey, C., Joussame, S., Kattsov, V., Kitoh, A., Ogana, W., Pitman, A.J., Weaver, A.J., Wood, R.A., Zhao, Z.-C., 2001. Model evaluation. In: Houghton, J.T., Ding, Y., Griggs, D.J., Noguer, M., van der Linden, P.J., Dai, X., Maskell, K., Johnson, C.A. (Eds.), *Climate Change 2001: The Scientific Basis. Contribution of Working Group I to the Third Assessment Report of the Intercontinental Panel on Climate Change*. Cambridge Univ. Press, UK, pp. 471–524.
- National Snow and Ice Data Center, 1996. Northern Hemisphere EASE-Grid Weekly Snow Cover and Sea Ice Extent. Volumes 1.0–2.0, EOSDIS NSIDC Distributed Active Archive Center (NSIDC DAAC). University of Colorado, Boulder.
- Robinson, D.A., Kukla, G., 1985. Maximum surface albedo of seasonally snow covered lands in the Northern Hemisphere. *J. Clim. Appl. Meteorol.* 24, 402–411.
- Samuelsson, P., Bringfelt, B., Graham, L.P., 2003. The role of aerodynamic roughness for the relationship between runoff and snow evaporation in land-surface schemes. *Glob. Planet. Change* 38, 93–99.
- Slater, A.G. et al., 2001. The representation of snow in land surface schemes: results from PILPS-2(d). *J. Hydrometeorol.* 2, 7–25.
- Takayabu, I., Takata, K., Yamazaki, T., Ueno, K., Yabuki, H., Haginoya, S., 2001. Comparison of the four land surface models driven by a common forcing data prepared from GAME/Tibet POP’97 products—snow accumulation and soil freezing processes. *J. Meteorol. Soc. Jpn.* 79, 535–554.
- van den Hurk, B., Viterbo, P., 2003. The Torne–Kalix PILPS2E experiment as a test bed for modifications to the ECMWF land surface scheme. *Glob. Planet. Change* 38, 165–173.
- Wood, E.F., Lettenmaier, D.P., Liang, X., Lohmann, D., Boone, A., Chang, S., Chen, F., Dai, Y., Dickinson, R., Duan, Q., Ek, M., Gusev, Y., Habets, F., Irannejad, P., Koster, R., Mitchell, K., Nasonova, O., Noilhan, J., Schaake, J., Schlosser, A., Shao, Y., Shmakin, A., Verseghy, D., Warrach, K., Wetzel, P., Xue, Y., Yang, Z., Zeng, Q., 1998. The project for intercomparison of land-surface parameterization schemes (PILPS) phase 2(c) Red-Arkansas River basin experiment: 1. Experiment description and summary intercomparisons. *Glob. Planet. Change* 19, 115–135.
- World Climate Research Programme (WCRP), 1999. Report on the hydrology models intercomparison planning meeting, Koblenz, Germany, 27–29 March 1999. WCRP Informal Report 12/1999.
- Xia, Y., Pitman, A.J., Leplastrier, M., Henderson-Sellers, A., 2003. The CHAMELEON Surface Model (CHASM): description and use with the PILPS Phase 2e forcing data. *Glob. Planet. Change* 38, 121–135.